SF 298 MASTER COPY KEEP THIS COPY FOR REPRODUCTION PURPOSES

REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 1997	Technical - 97-13

4. TITLE AND SUBTITLE
The Law of the Iterated Logarithm and Central Limit Theorem for L-Statistics

5. FUNDING NUMBERS
DAAH04-96-1-0082

6. AUTHOR(S)
Deli Li, M.B. Rao and R.J. Tomkins

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)
Center for Multivariate Analysis
417 Thomas Building
Penn State University
University Park, pA 16802

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER
ARO 35518.19-MA

11. SUPPLEMENTARY NOTES
The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)
The main idea in this paper is that we devise an effective way of combining the Smirnov's law of the iterated logarithm for empirical processes, and some well-known results of limit behavior of L-statistics to establish new results on the central limit theorem, law of the iterated logarithm, and strong law of large numbers, for L-statistics. We show further that this approach can be pursued profitably to obtain necessary and sufficient conditions for either almost sure convergence or convergence in distribution of some well-known L-statistics and U-statistics. A law of the logarithm for weighted sums of order statistics is stated with no proof.

19970820 015

14. SUBJECT TERMS
Central limit theorem; empirical processm laws of the iterated logarithm, L-statistics, strong law of large numbers.

15. NUMBER IF PAGES
32

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

THE LAW OF THE ITERATED LOGARITHM AND CENTRAL LIMIT THEOREM FOR L-STATISTICS

Deli Li, M.B. Rao and R.J. Tomkins

Technical Report 97-13

July 1997

Center for Multivariate Analysis
417 Thomas Building
Penn State University
University Park, PA 16802

Research supported by the Army Research Office under Grant DAAH04-96-1-0082.

THE LAW OF THE ITERATED LOGARITHM AND CENTRAL LIMIT THEOREM FOR L-STATISTICS

DELI LI[1], M. BHASKARA RAO[2] AND R. J. TOMKINS[3 4]

Abstract The main idea in this paper is that we devise an effective way of combining the Smirnov's law of the iterated logarithm for empirical processes, and some well-known results of limit behavior of L-statistics to establish new results on the central limit theorem, law of the iterated logarithm, and strong law of large numbers, for L-statistics. We show further that this approach can be pursued profitably to obtain necessary and sufficient conditions for either almost sure convergence or convergence in distribution of some well-known L-statistics and U-statistics. A law of the logarithm for weighted sums of order statistics is stated with no proof.

AMS 1991 subject classifications. Primary 60F15, 60F05; Secondary 62G30

[1]DEPARTMENT OF STATISTICS AND ACTUARIAL SCIENCES, THE UNIVERSITY OF WESTERN ONTARIO, LONDON, ONTARIO, CANADA N6A 5B7

[2]DEPARTMENT OF STATISTICS, NORTH DAKOTA STATE UNIVERSITY, FARGO, ND 58105, USA

[3]DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF REGINA, REGINA, SASKATCHEWAN, CANADA S4S 0A2

[4]RESEARCH SUPPORTED BY A GRANT FROM THE NATURAL SCIENCES AND ENGINEERING RESEARCH COUNCIL OF CANADA.

Keywords and Phrases: Central limit theorem, empirical process, laws of the iterated logarithm, L-statistics, strong law of large numbers.

1 Introduction

Throughout this article, $\{X, X_n;\ n \geq 1\}$ will denote a sequence of independent identically distributed real random variables (i.i.d. random variables) with common distribution function F given by $F(x) = P(X \leq x),\ x \in \mathcal{R}$, the real line. For each positive integer n, let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ be the order statistics of $X_1,\ X_2,\ \cdots,\ X_n$. Let H be a real valued measurable function defined on $\mathcal{R}$. Linear combinations of order statistics (in short, L-statistics) are statistics of the form

$$L_n = \frac{1}{n}\sum_{i=1}^{n} c_{i,n} H(X_{i:n})$$

where the weights $c_{i,n},\ 1 \leq i \leq n,\ n \geq 1$, are real numbers.

Research in L-statistics has been active for more than 20 years. Much is known about their limit behaviours including asymptotic normality, Berry-Esseen type bounds for normal approximation, Cramer type large deviations, the law of the iterated logarithm and the Kolmogorov type strong laws of large numbers under quite general conditions. See Helmers (1977), Mason (1982), Mason and Shorack (1992), Sen (1978), Shorack (1972), Stigler (1974), Van Zwet (1980), Wellner (1977a, b), etc and, in particular, two books by Serfling (1980) and Shorack and Wellner (1986) and references therein. To our knowledge, however, all results established for L-statistics need $H(\cdot)$ to be a known function of the form $H(\cdot) = H_1(\cdot) - H_2(\cdot)$ with each $H_i \uparrow$ and left continuous.

Furthermore, $H_1(\cdot)$ and $H_2(\cdot)$ should also satisfy

$$|H_i(F^{\leftarrow}(t))| \leq M_1 t^{-d_1}(1-t)^{-d_2}, \qquad 0 < t < 1 \tag{1.1}$$

with some fixed M_1, d_1 and d_2, where

$$F^{\leftarrow}(t) = \inf\{s;\ F(s) \geq t\}, \quad 0 < t < 1.$$

Some conditions on $\{c_{i,n};\ 1 \leq i \leq n,\ n \geq 1\}$ are also needed. One often chooses

$$c_{i,n} = J(\frac{i}{n}), \quad 1 \leq i \leq n,\ n \geq 1$$

where $J(t)$ satisfies some continuous condition in $(0,1)$ as well as

$$|J(t))| \leq M_2 t^{-b_1}(1-t)^{-b_2}, \qquad 0 < t < 1 \tag{1.2}$$

with some fixed M_2, b_1 and b_2. For obtaining either central limit theorem or the law of the iterated logarithm for

$$\sum_{i=1}^{n} J\left(\frac{i}{n}\right) H(X_{i:n}), \quad n \geq 1,$$

further condition on b_1, b_2, d_1 and d_2 is

$$a = \max\{b_1 + d_1, b_2 + d_2\} < \frac{1}{2}.$$

These conditions sometimes are not easy to be vertified.

<u>**Example 1.1**</u> Let $p > 1$ and $\{X,\ X_n;\ n \geq 1\}$ a sequence of i.i.d. random variables with common density function

$$f(x) = \begin{cases} fracC_p|x|^3(L|x|)^p, & |x| \geq e \\ 0, & |x| < e \end{cases}$$

where $C_p > 0$ is a constant such that $\int_{-\infty}^{\infty} f(x)dx = 1$. Here and the following $Lx = \log_e \max\{e, x\}$ and $L_2x = L(Lx)$, $x \in \mathcal{R}$. Clearly,

$$E(X) = 0 \quad \text{and} \quad E(X^2) = \frac{2C_p}{p-1}.$$

It is easy to check that

$$F(x) = P(X \leq x) \sim \begin{cases} 2^{-1}C_p x^{-2}(L|x|)^{-p} & \text{as } x \to -\infty, \\ 1 - 2^{-1}C_p x^{-2}(L|x|)^{-p} & \text{as } x \to \infty. \end{cases}$$

Hence

$$F^{\leftarrow}(t) \sim \begin{cases} \left(\frac{2}{C_p}\right)^{1/2} t^{-1/2}\left(L(\frac{1}{t})\right)^{-p} & \text{as } t \to 0^+, \\ \left(\frac{2}{C_p}\right)^{1/2} (1-t)^{-1/2}\left(L(\frac{1}{1-t})\right)^{-p} & \text{as } t \to 0^-. \end{cases}$$

Choose $H(x) = x$ and $J(t) \equiv 1$. It follows that one can not find b_1, b_2, d_1 and d_2 such that (1.1) and (1.2) hold and

$$a = \max\{b_1 + d_1,\ b_2 + d_2\} < \frac{1}{2}.$$

Thus, by using any result of the law of the iterated logarithm for L-statistics established so far, one can not say anything about

$$\limsup_{n\to\infty} \frac{\sum_{i=1}^{n} J(\frac{i}{n})X_{i:n}}{\sqrt{2n\log\log n}} = \limsup_{n\to\infty} \frac{\sum_{i=1}^{n} X_{i:n}}{\sqrt{2n\log\log n}}$$

which the classical Hartman-Wintner law of the iterated logarithm asserts that the limit superior equals to $(2C_p/(p-1))^{1/2}$ almost surely. As Wellner (1977b) mentioned, the classical law of the iterated logarithm does follow only if the second moment condition is strengthened to r-th moment condition for some $r > 2$. See Example 1a on page 492 in Wellner (1977b).

Example 1.2 Let $\mathbf{W}$ be a standard Wiener process on $\mathcal{R}$ (i.e. $\{\mathbf{W}(t);\ t \geq 0\}$ and $\{\mathbf{W}(-t);\ t \geq 0\}$ are two independent copies of Wiener process starting from 0), and let $\{X, X_n;\ n \geq 1\}$ be a sequence of i.i.d. random variables, independent of $\mathbf{W}$. For any given path of $\mathbf{W}$, we consider L-statistics

$$\sum_{i=1}^{n} \frac{i}{n} \mathbf{W}(X_{i:n}), \quad n \geq 1.$$

As we know, for almost every path of $\mathbf{W}$, $\mathbf{W}$ can not be represented as the form $\mathbf{W} = \mathbf{W}_1 - \mathbf{W}_2$ with $\mathbf{W}_i \uparrow$ and left continuous. One also can not say anything about

$$\limsup_{n \to \infty} \frac{\sum_{i=1}^{n} \frac{i}{n} \mathbf{W}(X_{i:n}) - A_n}{\sqrt{2n \log \log n}}$$

for some sequence $\{A_n;\ n \geq 1\}$, by using any result of the law of the iterated logarithm for L-statistics established so far.

Motivated by these examples above, one of the main objectives of this paper is to find necessary and sufficient conditions for the law of the iterated logarithm and the central limit theorem, respectively, of L-statistics under some very relaxed conditions on $H(\cdot)$. In fact, from Theorems 2.1 and 2.2 in this paper, we can find that the only condition on $H(\cdot)$ we need is $E(|H(X)|) < \infty$. We should note that $J(\cdot)$ is strengthened to be a Lipschitz function of order 1 when $H(\cdot)$ is relaxed.

It is valuable to give an outline of the main idea of proofs of Theorems 2.1 and 2.2. Throughout this paper, $\{U, U_n;\ n \geq 1\}$ represents a sequence of *iid*

random variables with uniform $(0,1)$ distribution. Then, we have

$$\{X,\ X_n;\ n \geq 1\} \stackrel{d}{=} \{F^{\leftarrow}(U),\ F^{\leftarrow}(U_n);\ n \geq 1\}$$

where "$\stackrel{d}{=}$" means equal in distribution. This is a well-known fact. It now follows that

$$\{X_{i:n};\ 1 \leq i \leq n,\ n \geq 1\} \stackrel{d}{=} \{F^{\leftarrow}(U_{i:n});\ 1 \leq i \leq n,\ n \geq 1\}$$

where $U_{i:n},\ 1 \leq i \leq n$, are the order statistics of $U_i,\ 1 \leq i \leq n$. Note that

$$P\left(U_i \neq U_j \ \text{ for all } \ 1 \leq < j < \infty\right) = 1.$$

So we have that

$$\begin{aligned}
\textstyle\sum_{i=1}^{n} J(\frac{i}{n}) H\left(F^{\leftarrow}(U_{i:n})\right) \quad &= \quad \textstyle\sum_{i=1}^{n} J(U_{i:n}) H\left(F^{\leftarrow}(U_{i:n})\right) \\
&\quad + \ \textstyle\sum_{i=1}^{n} \left(J(\frac{i}{n}) - J(U_{i:n})\right) H\left(F^{\leftarrow}(U_{i:n})\right) \\
&\stackrel{\text{a.s.}}{=} \quad \textstyle\sum_{i=1}^{n} J(U_i) H\left(F^{\leftarrow}(U_i)\right) \\
&\quad + \ \textstyle\sum_{i=1}^{n} \left(J(\mathcal{D}_n(U_{i:n})) - J(U_{i:n})\right) H\left(F^{\leftarrow}(U_{i:n})\right) \\
&\triangleq \quad S_n + R_n, \qquad n \geq 1
\end{aligned}$$

where $\mathcal{D}_n$ is the empirical distribution function of $U_1,\ U_2, \ldots,\ U_n$. Clearly, classical results can be applied to the first part S_n above. As for the second part R_n, we may try to apply some known results for the empirical processes to R_n to obtain what we want. However, we can not misunderstand that, under some conditions (except $J(t) \equiv C$ some constant), both the law of the iterated logarithm and the central limit theorem for $\sum_{i=1}^{n} J(\frac{i}{n}) H(X_{i:n}),\ n \geq 1$, only

depend on $\sum_{i=1}^{n} J(U_i)H(F^{\leftarrow}(U_i))$, $n \geq 1$. We shall know this from Theorems 2.1 and 2.2.

This approach is quite simple. The empirical process is a powerful tool which has now become a standard technique in proving limit theorems. See, for example, Gilat and Hill (1992) Mason and Shorack (1992), Shorack and Wellner (1986) and references therein. In Section 3, we show further that this approach can be applied to obtain necessary and sufficient conditions for either almost sure convergence or convergence in distribution of some well-known L-statistics or U-statistics. This is the second main objective of this paper.

It is very natural to consider the case of general scores $c_{i,n}$, $1 \leq i \leq n$, $n \geq 1$ with

$$\sup_{1 \leq i \leq n, n \geq 1} |c_{i,n}| < \infty.$$

In this connection, one may ask what we can say about the law of the iterated logarithm for

$$\sum_{i=1}^{n} c_{i,n} H(X_{i:n}), \quad n \geq 1.$$

We shall have some certain sense about this problem by a result on what we call a law of the logarithm for L-statistics. See Section 4.

2 Main Results

In this section, we give our main results of this paper as well as their proofs.

Let X be a real random variable with distribution function $F(x)$ and U

a random variable with uniform $(0,1)$ distribution. Let $H(\cdot)$ be a real Borel-measurable function defined on $\mathcal{R}$ with

$$E(|H(X)|) < \infty. \tag{2.1}$$

Then we write

$$\mu = \mu(F, J, H) = E\left(J(U)H(F^{\leftarrow}(U)\right)$$

exists for all $J(\cdot) \in C[0,1]$, the set of continuous functions defined on $[0,1]$. For every given Lipschitz function $J(\cdot)$ of order 1 defined on $[0,1]$, write

$$\begin{cases} Z = J(U)H\left(F^{\leftarrow}(U)\right), \\ Y = -Z + \mu - \int_0^1 \left(I_{\{U \le t\}} - t\right) J'(t)H\left(F^{\leftarrow}(t)\right) dt. \end{cases} \tag{2.2}$$

Then Y and Z are two random variables under (2.1). It is easy to see that

$$E(|Y|) < \infty \text{ and } E(Y) = 0 \tag{2.3}$$

and that

$$\sigma^2 = \mathrm{Var}(Y) = E(Y^2) < \infty \tag{2.4}$$

if and only if

$$E(Z^2) < \infty. \tag{2.5}$$

Let $\{\xi_n;\ n \ge 1\}$ be a sequence of random variables. We say that $\{\xi_n;\ n \ge 1\}$ is bounded in probability if

$$\lim_{x\to\infty} \sup_{n \ge 1} P(|\xi_n| \ge x) = 0.$$

The main results in this paper are following Theorems 2.1 and 2.2.

Theorem 2.1 (Necessary and Sufficient Conditions for the LIL of L-Statistics) Let $\{X, X_n;\ n \geq 1\}$ be a sequence of i.i.d. random variables and $H(\cdot)$ a real Borel-measurable function defined on $\mathcal{R}$ such that (2.1) holds. Let $J(\cdot)$ be a Lipschitz function of order 1 defined on $[0,1]$. Then

$$\limsup_{n\to\infty} \frac{\left|\sum_{i=1}^{n} J(\frac{i}{n})H(X_{i:n}) - n\mu\right|}{\sqrt{2n\log\log n}} < \infty \quad \text{a.s.} \tag{2.6}$$

if and only if (2.5) holds. In either case,

$$\limsup_{n\to\infty}(\liminf_{n\to\infty}) \frac{\sum_{i=1}^{n} J(\frac{i}{n})H(X_{i:n}) - n\mu}{\sqrt{2n\log\log n}} \overset{+}{=}(-)\,\sigma \quad \text{a.s.} \tag{2.7}$$

where σ^2 is defined by (2.4).

Theorem 2.2 (Necessary and Sufficient Conditions for the CLT of L-Statistics)

Under the conditions of Theorem 2.1, we have that

$$\left\{ \frac{\sum_{i=1}^{n} J(\frac{i}{n})H(X_{i:n}) - n\mu}{\sqrt{n}};\ n \geq 1 \right\} \text{ is bounded in probability} \tag{2.8}$$

if and only if (2.5) holds. In either case,

$$\frac{\sum_{i=1}^{n} J(\frac{i}{n})H(X_{i:n}) - n\mu}{\sqrt{n}} \xrightarrow{d} N(0,\sigma^2) \tag{2.9}$$

where "$\xrightarrow{d}$" means convergence in distribution.

Two remarks are in order.

(i) Theorems 2.1 and 2.2, respectively, include the classical LIL and the classical CLT, respectively, as special cases. (Take $J(t) = 1, t \in [0,1]$ and $H(x) = x,\ x \in \mathcal{R}$.)

(ii) From the proofs of Theorems 2.1 and 2.2, one can replace $J(\frac{i}{n})$, $1 \le i \le n$, $n \ge 1$ by $J(t_{i:n})$, $1 \le i \le n$, $n \ge 1$ with $t_{i,n} \in [0,1]$, $1 \le i \le n$ and $\max_{1\le i\le n} \left|t_{i,n} - \frac{i}{n}\right| = o(\frac{1}{\sqrt{n}})$ as $n \to \infty$.

<u>**Proof of Theorem 2.1**</u> Note that (1..1), (1.2) and (1.3) hold and that the classical LIL asserts that

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{|S_n - n\mu|}{\sqrt{2n\log\log n}} \\
&= \limsup_{n\to\infty} \frac{\left|\sum_{i=1}^{n} (J(U_i)H(F^{\leftarrow}(U_i)) - \mu)\right|}{\sqrt{2n\log\log n}} \qquad (2.10) \\
&< \infty \quad \text{a.s.}
\end{aligned}$$

if and only if (2.4) holds. So the first part of the theorem follows provided we show that, under the conditions of the theorem,

$$\limsup_{n\to\infty} \frac{|R_n|}{\sqrt{2n\log\log n}} \le 2^{-1}B(J)E(|H(X)|) \quad \text{a.s.} \qquad (2.11)$$

where

$$B(J) = \sup_{0\le s<t\le 1} \left|\frac{J(s) - J(t)}{s - t}\right| < \infty.$$

Clearly,

$$\begin{aligned}
\frac{|R_n|}{\sqrt{2n\log\log n}} &\le \max_{1\le i\le n} |J(\mathcal{D}_n(U_{i:n})) - J(U_{i:n})| \cdot \frac{\sum_{i=1}^{n} |H(F^{\leftarrow}(U_{i:n}))|}{\sqrt{2n\log\log n}} \\
&\le B(J)\left(\max_{1\le i\le n} |\mathcal{D}_n(U_{i:n}) - U_{i:n}|\right) \cdot \frac{\sum_{i=1}^{n} |H(F^{\leftarrow}(U_i))|}{\sqrt{2n\log\log n}} \\
&\le B(J)\left(\frac{n}{2\log\log n}\right)^{1/2} \sup_{0\le t\le 1} |\mathcal{D}_n(t) - t| \cdot \left(\frac{1}{n}\right)\sum_{i=1}^{n} |H(F^{\leftarrow}(U_i))|.
\end{aligned}$$

By Smirnov's (1944) law of the iterated logarithm,

$$\limsup_{n\to\infty} \left(\frac{n}{2n\log\log n}\right)^{1/2} \sup_{0\le t\le 1} |\mathcal{D}_n(t) - t| = \frac{1}{2} \quad \text{a.s.}$$

and by Kolmogorov's strong law of large numbers,

$$\lim_{n\to\infty}\left(\frac{1}{n}\right)\sum_{i=1}^{n}|H(F^{\leftarrow}(U_i))| = E(|H(F^{\leftarrow}(U))|) = E(|H(X)|) \quad \text{a.s.}$$

Consequently, (2.11) and hence the first part of the theorem follows.

We now prove that conditions (2.1) and (2.5) imply that (2.7) holds. Consider the following two Banach spaces

$$\mathbf{B}_1 = \left\{h(\cdot);\ \|h(\cdot)\|_1 = \int_0^1 |h(F^{\leftarrow}(t))|dt < \infty\right\}$$

and

$$\mathbf{B}_2 = \left\{h(\cdot);\ \|h(\cdot)\|_2 = \left(\int_0^1 J^2(t)h^2(F^{\leftarrow}(t))dt\right)^{1/2}\right\}.$$

Let $\mathcal{M}$ denote the class of all real valued functions $h(\cdot)$ defined on $\mathcal{R}$ satisfying

$$h'(x) \text{ is continuous on } \mathcal{R}$$

and

$$h'(x) = 0 \text{ when } |x| \text{ is large enough.}$$

It is easy to show that $\Updownarrow$ is a dense subset of both $\mathbf{B}_1$ and $\mathbf{B}_2$. It follows that, for any given $\varepsilon > 0$, since (2.1) and (2.5) hold, there exists $H_\varepsilon(\cdot) \in \mathcal{M}$ such that

$$E\left(|H(X) - H_\varepsilon(X)|\right) \le \varepsilon \qquad \text{Var}(Y - Y_\varepsilon) \le \varepsilon$$

and

$$\text{Var}(Z - Z_\varepsilon) \le \varepsilon$$

where

$$Z_\varepsilon = J(U)H_\varepsilon(F^{\leftarrow}(U)) \qquad \mu_\varepsilon = E(Z_\varepsilon)$$

$$Y_\varepsilon = -Z_\varepsilon + \mu_\varepsilon - \int_0^1 (I_{\{U\le t\}} - t)J'(t)H_\varepsilon(F^{\leftarrow}(t))dt.$$

Write

$$\hat{H}_\varepsilon(\cdot) = H(\cdot) - H_\varepsilon(\cdot) \quad \text{and} \quad \hat{\mu}_\varepsilon = \mu - \mu_\varepsilon.$$

Noting (1.3), (2.10) and (2.11), we have that

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{\left|\sum_{i=1}^n J(\frac{i}{n})\hat{H}_\varepsilon(X_{i:n}) - n\hat{\mu}_\varepsilon\right|}{\sqrt{2n\log\log n}} \\
&\le \sqrt{\operatorname{Var}(Z - Z_\varepsilon)} + 2^{-1}B(J)E(|\hat{H}_\varepsilon(X)|) \qquad (2.12)\\
&\le \varepsilon^{1/2} + \tfrac{1}{2}B(J)\varepsilon \quad \text{a.s.}
\end{aligned}$$

Let

$$\begin{aligned}
H_\varepsilon^+(x) &= \int_{-\infty}^x H_\varepsilon'(x)I_{\{H_\varepsilon'(x)\ge 0\}}dx,\\
H_\varepsilon^-(x) &= -\int_{-\infty}^x H_\varepsilon'(x)I_{\{H_\varepsilon'(x)<0\}}dx.
\end{aligned}$$

Then

$$H_\varepsilon(x) = H_\varepsilon^+(x) - H_\varepsilon^-(x).$$

One now can easily check that $J(\cdot)$ and $H_\varepsilon(\cdot)$ satisfy all conditions for Theorem 4 of Wellner (1977b) (See also Theorem 19.2 on page 665 of Shorack and Wellner (1986)). We thus have

$$\limsup_{n\to\infty}(\liminf_{n\to\infty}) \frac{\sum_{i=1}^n J(\frac{i}{n})H_\varepsilon(X_{i:n}) - n\mu_\varepsilon}{\sqrt{2n\log\log n}} \overset{+}{=}(-)\,\sigma_\varepsilon \quad \text{a.s.} \qquad (2.13)$$

where

$$\sigma_\varepsilon^2 = \operatorname{Var}\left(\int_0^1 (I_{\{U\le t\}} - t)J(t)dH_\varepsilon(F^{\leftarrow}(t))\right).$$

Since $J(\cdot)$ is a Lipschitz function of order 1 defined on $[0,1]$, we have that

$$\begin{aligned}&\int_0^1 (I_{\{U\le t\}} - t)J(t)dH_\varepsilon(F^{\leftarrow}(t))\\ &\overset{\text{a.s.}}{=} -J(U)H_\varepsilon(F^{\leftarrow}(U)) + \mu_\varepsilon - \int_0^1 (I_{\{U\le t\}} - t)J'(t)H_\varepsilon(F^{\leftarrow}(t))dt\\ &= Y_\varepsilon.\end{aligned}$$

Thus

$$\sigma_\varepsilon^2 \longrightarrow \sigma^2 = \mathrm{Var}(Y) \quad \text{as } \varepsilon \downarrow 0.$$

Letting $\varepsilon \downarrow 0$ and combining (2.12) with (2.13), (2.7) follows, and the proof of the theorem is complete. $\square$

<u>**Proof of Theorem 2.2**</u> A proof of Theorem 2.2 can be culled from the proof of Theorem 2.1 with obvious modifications. In fact, instead of using Smirnov's (1944) LIL for empirical processes, we use the following celebrated result which is due to Kolmogrov (1933). That is, for all $\lambda > 0$,

$$\lim_{n\to\infty} P\left(\sqrt{n}\sup_{0\le t\le 1}|\mathcal{D}_n(t) - | \ge \lambda\right) = 2\sum_{k=0}^{\infty}\exp(-2k^2\lambda^2).$$

Thus, Kolmogrov's SLLN implies that (2.8) holds if and only if

$$\left\{\frac{\sum_{i=1}^n J(U_i)H(F^{\leftarrow}(U_i)) - n\mu}{\sqrt{n}};\ n\ge 1\right\} \quad \text{is bounded in probability}$$

which is equivalent to (2.5). Thus, the first part of Theorem 2.2 follows.

As for the second part of Theorem 2.2, note that, for any $\delta > 0$.

$$\limsup_{\varepsilon\downarrow 0}\limsup_{n\to\infty} P\left(\frac{|\sum_{i=1}^n J(\frac{i}{n})(H(X_{i:n}) - H_\varepsilon(X_{i:n})) - n(\mu-\mu_\varepsilon)|}{\sqrt{n}} \ge \delta\right) = 0 \tag{2.14}$$

and by Theorem 19.1 on page 664 of Shorack and Wellner (1985),

$$\frac{\sum_{i=1}^{n} J(\frac{i}{n})H_{\varepsilon}(X_{i:n}) - n\mu_{\varepsilon}}{\sqrt{n}} \xrightarrow{\text{d}} N(0, \sigma_{\varepsilon}^{2}). \tag{2.15}$$

Combining (2.14) with (2.15), (2.9) follows. □

3 Applications and Examples

In this section, weshow further how either our approach addressed in Section 1 or Theorems 2.1 and 2.2 can be applied to obtain necessary and sufficient conditions for either almost sure convergence or convergence in distribution of some L-statistics or even U-statistics.

As an application of our approach, we state the following interesting result. Its proof is left to the readers.

<u>**Theorem 3.1**</u> Let $\{X, X_n;\ n \geq 1\}$ be a sequence of i.i.d. random variables and $H(\cdot)$ a real Borel-measurable function defined on $\mathcal{R}$. Let $J(\cdot)$ be a Lipschitz function of order 1 defined on $[0, 1]$. Then, we have

(I) If there exists a real sequence $\{b_n;\ n \geq 1\}$ such that

$$b_{n \longrightarrow \infty} \quad \text{as } n \longrightarrow \infty \tag{3.1}$$

and

$$\lim_{n \to \infty} \frac{\sum_{i=1}^{n} |H(X_i)|}{n\sqrt{b_n}} = 0 \quad \text{a.s.} \tag{3.2}$$

then, for any real sequence $\{A_n;\ n \geq 1\}$,

$$\begin{aligned}
&\limsup_{n\to\infty}(\liminf_{n\to\infty})\frac{\sum_{i=1}^{n} J(\frac{i}{n})H(X_{i:n})-A_n}{\sqrt{2n(\log\log n)b_n}} \\
&= \limsup_{n\to\infty}(\liminf_{n\to\infty})\frac{\sum_{i=1}^{n} J(U_i)H(F^{\leftarrow}(U_i))-A_n}{\sqrt{2n(\log\log n)b_n}} \quad \text{a.s.}
\end{aligned} \tag{3.3}$$

(ii) If there exist sequence $\{b_n;\ n \geq 1\}$ and $\{A_n;\ n \geq 1\}$ such that (3.2) holds and

$$\frac{\sum_{i=1}^{n} J(U_i)H(F^{\leftarrow}(U_i)) - A_n}{\sqrt{nb_n}} \xrightarrow{\mathrm{d}} G(x) \tag{3.4}$$

where $G(x)$ is a distribution function, then

$$\frac{\sum_{i=1}^{n} J(\frac{i}{n})H(X_{i:n}) - A_n}{\sqrt{nb_n}} \xrightarrow{\mathrm{d}} G(x) \tag{3.5}$$

As a corolaary of Theorem 3.1, we have

<u>**Corollary 3.2**</u> Let $H(\cdot)$ and $J(\cdot)$ be the same as in Theorem 3.1. Let

$$Z = J(U)H(F^{\leftarrow}(U)).$$

If

$$E(|H(X)|) < \infty, \qquad E(Z^2) = \infty$$

and Z is in the domain of attraction of the normal distribution, then, there exist sequence $\{A_n;\ n \geq 1\}$ and $\{B_n > 0;\ n \geq 1\}$ such that

$$\frac{\sum_{i=1}^{n} J(\frac{i}{n})H(X_{i:n}) - A_n}{B_n} \xrightarrow{\mathrm{d}} N(0,1) \tag{3.6}$$

where B_n may be chosen as

$$B_n = \sup\left\{c:\ c^{-2}E(Z^2 I_{\{|Z|<c\}}) \geq \frac{1}{n}\right\}$$

and A_n may be taken as

$$A_n = \frac{n}{B_n} E(ZI\{|Z| < B_n\}).$$

Proof Since Z is the domain of attraction of the normal distribution, we have that

$$\frac{\sum_{i=1}^{n} Z_i - A_n}{B_n} \xrightarrow{\text{d}} N(0,1)$$

for some $\{A_n;\ n \geq 1\}$ and $\{B_n;\ n \geq 1\}$ and they may be chosen as above, where $\{Z_n;\ n \geq 1\}$ is a sequence of i.i.d. random variables with common distribution function as Z's. Note that $E(Z^2) = \infty$. So

$$B_n = \sqrt{n(B^2/n)} = \sqrt{nb_n}, \qquad n \geq 1$$

with $b_n = b_n^2/n \longrightarrow \infty$ as $n \longrightarrow \infty$. Now, it follows that $E(|H(X)|) < \infty$ implies that (3.2) holds. By Theorem 3.1, (3.6) follows. □

<u>**Remark**</u> It is interesting to note that, when $E(|H(X)|) < \infty$, $E(Z^2) = \infty$ and Z is in the domain of attraction of the normal distribution, A_n and B_n are determined by Z, not Y.

Let $\{X, X_n;\ n \geq 1\}$ be a sequence of i.i.d. random variables. Gini's mean difference,

$$\frac{2}{n(n-1)} \sum_{1 \leq i \leq j \leq n} |X_i - X_j|$$

considered as a U-statistic for unbiased estimation of the dispersion parameter

$$\theta = E(|X_1 - X_2|),$$

is introduced in both Serfling (1980, page 263) and Shorack and Wellner (1986, page 676). It may be represented as an L-statistic as follows

$$\frac{2}{n(n-1)} \sum_{1 \le i \le j \le n} |X_i - X_j| = \frac{1}{n} \sum_{i=1}^{n} (4 \cdot \frac{i-1}{n-1} - 2) X_{i:n}. \tag{3.7}$$

Using Theorems 2.1 and 2.2 and our approach, we can establish the following result.

<u>**Theorem 3.3**</u> **(i)** (Kolmogorov's SLLN for Gini's mean difference)

$$\limsup_{n \to \infty} \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} |X_i - X_j| < \infty \quad \text{a.s.} \tag{3.8}$$

if and only if

$$E(|X|) < \infty. \tag{3.9}$$

In either case,

$$\lim_{n \to \infty} \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} |X_i - X_j| = \theta = E(|X_1 - X_2|) \quad \text{a.s.} \tag{3.10}$$

(ii) (LIL for Gini's mean difference) For some θ,

$$\limsup_{n \to \infty} \sqrt{\frac{n}{2 \log \log n}} \left| \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} |X_i - X_j| - \theta \right| < \infty \quad \text{a.s.} \tag{3.11}$$

if and only if

$$E(X^2) < \infty. \tag{3.12}$$

In either case, $\theta = E(|X_1 - X_2|)$ and

$$\limsup_{n \to \infty} (\liminf_{n \to \infty}) \sqrt{\frac{n}{2 \log \log n}} \left(\frac{2}{n(n-1)} \sum_{1 \le i < j \le n} |X_i - X_j| - \theta \right) = \overset{+}{(-)} \sigma \quad \text{a.s.} \tag{3.13}$$

where

$$\sigma^2 = \mathrm{Var}(Y)$$
$$= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty}|y-x|dF(x)\right]^2 dF(y) - \theta^2 \tag{3.14}$$

and

$$Y = -(2U-1)F^{\leftarrow}(U) + \theta - 2\int_0^1 (I_{\{U\le t\}} - t)F^{\leftarrow}(t)dt.$$

(iii) (CLT for Gini's mean difference) For some θ,

$$\left\{\sqrt{n}\left(\frac{2}{n(n-1)}\sum_{1\le i<j\le n}|X_i - X_j| - \theta\right);\ n\ge 1\right\} \text{ is bounded in probability} \tag{3.15}$$

if and only if (3.12) holds. In either case, $\theta = E(|X_1 - X_2|)$

$$\sqrt{n}\left(\frac{2}{n(n-1)}\sum_{1\le i<j\le n}|X_i - X_j| - \theta\right) \xrightarrow{\mathrm{d}} N(0,\sigma^2) \tag{3.16}$$

where σ^2 is defined as in (3.14).

Remark Shorack and Wellner (1986, page 677) obtained (3.16) under the assumption that $E(|X|^{2+\delta}) < \infty$ for some $\delta > 0$.

Proof of Theorem 3.3 Take $J(t) = 4t-2,\ 0\le t\le 1$ and $H(x) = x,\ x\in \mathcal{R}$. Then (3.9) implies that

$$\lim_{n\to\infty}\frac{\sum_{i=1}^n J(U_i)F^{\leftarrow}(U_i)}{n} = \int_0^1 (4t-2)F^{\leftarrow}(t)dt$$
$$= \int_0^1\int_0^1 |F^{\leftarrow}(t) - F^{\leftarrow}(s)|\, dsdt$$
$$= \theta \quad \text{a.s.}$$

and

$$\lim_{n\to\infty}\frac{\sum_{i=1}^n |X_i|}{n\sqrt{b_n}} = 0 \quad \text{a.s.} \tag{3.17}$$

where $b_n = \sqrt{n/\log\log n} \longrightarrow \infty$ as $n \longrightarrow \infty$. By Theorem 3.1, we have that

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} J(\frac{i}{n})X_{i:n}}{n} = 0 \quad \text{a.s.} \tag{3.18}$$

Note that (3.9) also implies that

$$\left| \frac{\sum_{i=1}^{n} \left(J(\frac{i-1}{n-1}) - J(\frac{i}{n})\right) X_{i:n}}{n} \right| \le \frac{4}{n^2}\sum_{i=1}^{n} |X_i| \longrightarrow 0 \quad \text{a.s.} \quad \text{as } n \longrightarrow \infty. \tag{3.19}$$

So (3.7), (3.18) and (3.19) together imply that (3.10) holds.

We now prove that (3.8) implies (3.9). Obviously, (3.8) implies that

$$\limsup_{n\to\infty} \frac{2}{2n(2n-1)} \sum_{i=1}^{n} |X_{2i-1} - X_{2i}| < \infty \quad \text{a.s.}$$

Since $\{|X_{2n-1} - X_{2n}|;\ n \ge 1\}$ is a sequence of i.i.d. random variables, it follows that $E(|X_1 - X_2|^{1/2}) < \infty$ which is equivalent to

$$E(|X_1|^{1/2}) < \infty. \tag{3.20}$$

Thus (3.19) holds. Combining (3.8) with (3.19), we have that

$$\limsup_{n\to\infty} \frac{|\sum_{i=1}^{n} J(\frac{i}{n})X_{i:n}|}{n} < \infty \quad \text{a.s.} \tag{3.21}$$

Note that (3.20) is equivalent to

$$\lim_{n\to\infty} \frac{\sum_{i=1}^{n} |X_i|}{n\sqrt{n^2}} = 0 \quad \text{a.s.}$$

By Theorem 3.1 and noting (3.20),

$$\begin{aligned} \limsup_{n\to\infty} \frac{|\sum_{i=1}^{n} J(U_i)F^{-}(U_i)|}{\sqrt{2n(\log\log n)n^2}} &= \limsup_{n\to\infty} \frac{|\sum_{i=1}^{n} J(\frac{i}{n})X_{i:n}|}{\sqrt{2n(\log\log n)n^2}} \\ &= 0 \quad \text{a.s.} \end{aligned}$$

which implies that, for all $\varepsilon > 0$,

$$E\left(|J(U)F^{\leftarrow}(U)|^{\frac{2}{3}-\varepsilon}\right) = \int_0^1 |(4t-2)F^{\leftarrow}(t)|^{\frac{2}{3}-\varepsilon} dt < \infty. \tag{3.22}$$

Since $|J(0)J(1)| = 4 > 0$, it is easy to check that (3.22) is equivalent to

$$\begin{aligned} \int_0^1 |F^{\leftarrow}(t)|^{\frac{2}{3}-\varepsilon} dt &= \int_{-\infty}^{\infty} |x|^{\frac{2}{3}-\varepsilon} dF(x) \\ &= E(|X|^{\frac{2}{3}-\varepsilon}) < \infty. \end{aligned}$$

In particular, we have that

$$E(|X|^{7/12}(L_2|X|)) < \infty$$

which is equivalent to

$$\lim_{n\to\infty} \frac{\sum_{i=1}^n |X_i|}{n\sqrt{n^{10/7}/\log\log n}} = 0 \quad \text{a.s.}$$

Using Theorem 3.1 and noting (3.21) again, we have that

$$\begin{aligned} \lim_{n\to\infty} \frac{|\sum_{i=1}^n J(U_i)F^{\leftarrow}(U_i)|}{\sqrt{2n(\log\log n)n^{10/7}/\log\log n}} &= \lim_{n\to\infty} \frac{|\sum_{i=1}^n J(U_i)F^{\leftarrow}(U_i)|}{\sqrt{2n^{17/7}}} \\ &= 0 \quad \text{a.s.} \end{aligned}$$

which is equivalent to

$$E(|J(U)F^{\leftarrow}(U)|^{14/17}) < \infty$$

and hence

$$E(|X|^{14/17}) < \infty. \tag{3.23}$$

Since $\frac{3}{2} > \frac{17}{14}$, (3.23) implies that

$$\lim_{n\to\infty} \frac{\sum_{i=1}^n |X_i|}{n\sqrt{n/\log\log n}} = 0 \quad \text{a.s.}$$

Using Theorem 3.1 and noting (3.21) again, we have that

$$\begin{aligned}\lim_{n\to\infty} \frac{|\sum_{i=1}^{n} J(U_i)F^{\leftarrow}(U_i)|}{\sqrt{2n(\log\log n)n/\log\log n}} &= \lim_{n\to\infty} \frac{|\sum_{i=1}^{n} J(U_i)F^{\leftarrow}(U_i)|}{\sqrt{2}n} \\ &< \infty \quad \text{a.s.}\end{aligned}$$

which is equivalent to

$$E(|J(U)F^{\leftarrow}(U)|) < \infty$$

and hence (3.9) follows. The proof of (i) therefore complete.

Using the same technique as it has just been used above, one can show that either case of (3.11) and (3.15) implies that (3.12) holds and $\theta = E(|X_1 - X_2|)$. Now applying Theorems 2.1 and 2.2, (ii) and (iii) follow. □

Another important application of our approach and results is on the expectation of order statistics. Let $\theta_k = E(X_{1:k})$ where $X_{1:k}$ is the minimum value (smallest order statistic) in a sample of size k. Then an estimate of θ_k based on a sample of size n is

$$T_n = \binom{n}{k}^{-1} \sum_{(n,k)} \min\{X_{i_1}, \cdots, X_{i_k}\}$$

where the sum $\sum_{(n,k)}$ is taken over all subsets $1 \leq i_1 < \cdots < i_k \leq n$ of $\{1, 2, ..., n\}$. Obviously, T_n is also a U-statistic.

As it is deduced in Lee (1990, page 65), T_n can also be represented as an

L-statistic as follows

$$\begin{aligned} T_n &= \binom{n}{k}^{-1} \sum_{i=1}^{n-k} \binom{n-i}{k-1} X_{i:n} \\ &= \frac{1}{\binom{n}{k}(k-1)!} \sum_{i=1}^{n} (n-i)(n-i-1)\cdots(n-i-k+2) X_{i:n} \\ &= \frac{1}{\binom{n}{k}(k-1)!} \left(\sum_{i=1}^{n} (n-i)^{k-1} X_{i:n} + C_1 \sum_{i=1}^{n} (n-i)^{k-2} X_{i:n} + \cdots + C_{k-2} \sum_{i=1}^{n} (n-i) X_{i:n} \right) \end{aligned}$$

where $C_1,\ C_2, ..., C_{k-2}$ are constants depending on k only. Note that

$$\binom{n}{k} \sim \frac{n^k}{k!} \quad \text{as } n \longrightarrow \infty$$

and, for $i = 1, 2, ..., n$,

$$\frac{(n-i)^{n-j}}{(n^k/k!)\,(k-1)!} = \begin{cases} \frac{k}{n}\left(1 - \frac{i}{n}\right)^{k-1}, & j = 1, \\ O\left(\frac{1}{n^2}\right), & j \geq 2. \end{cases}$$

So we take $J(t) = (1-t)^{k-1},\ 0 \leq t \leq 1$ and $H(x) = x,\ x \in \mathcal{R}$. Using Theorems 3.1, 2.1 and 2.2, we have the following results.

Theorem 3.4 Let $\{X, X_n;\ n \geq 1\}$ be a sequence of i.i.d. random variables. Then, we have

(1) If there exists $p > \frac{2}{3}$ such that

$$E(|X|^p) < \infty \tag{3.24}$$

then the following three statements are equivalent

$$E\left(|\min\{X_1, X_2, ..., X_k\}|\right) < \infty \tag{3.25}$$

$$\limsup_{n\to\infty} \frac{k}{n} \left| \sum_{i=1}^{n} (1 - \frac{i}{n})^{k-1} X_{i:n} \right| < \infty \quad \text{a.s.} \tag{3.26}$$

$$\limsup_{n\to\infty} |T_n| < \infty \quad \text{a.s.} \tag{3.27}$$

In every case, we have that

$$\lim_{n\to\infty} T_n = \lim_{n\to\infty} \frac{k}{n}\sum_{i=1}^{n}\left(1-\frac{i}{n}\right)^{k-1} X_{i:n} = \theta_k \quad \text{a.s.} \tag{3.28}$$

(ii) If $E(|X|) < \infty$, then the following five statements are equivalent:

$$E\left((1-U)^{2k-2}(F^{\leftarrow}(U))^2\right) = \int_{-\infty}^{\infty}(1-F(x))^{2k-2}x^2 dF(x) < \infty \tag{3.29}$$

$$\limsup_{n\to\infty}\left(\frac{n}{2\log\log n}\right)^{1/2}|T_n-\theta_k| < \infty \quad \text{a.s.} \tag{3.30}$$

$$\limsup_{n\to\infty}\frac{k}{sqrt2n\log\log n}\left|\sum_{i=1}^{n}\left(1-\frac{i}{n}\right)^{k-1}X_{i:n}-\frac{\theta_k}{k}\right| < \infty \quad \text{a.s.} \tag{3.31}$$

$$\left\{n^{1/2}(T_n-\theta_k);\ n\geq 1\right\} \text{ is bounded in probability,} \tag{3.32}$$

$$\left\{\frac{k}{\sqrt{n}}\left(\sum_{i=1}^{n}(1-\frac{i}{n})^{k-1}X_{i:n}-\frac{\theta_k}{k}\right);\ n\geq 1\right\} \text{ is bounded in probability.} \tag{3.33}$$

In every case, we have that

$$\begin{aligned}
&\limsup_{n\to\infty}(\liminf_{n\to\infty})\left(\frac{n}{2\log\log n}\right)^{1/2}(T_n-\theta_k)\\
&=\limsup_{n\to\infty}(\liminf_{n\to\infty})\frac{k}{\sqrt{2n\log\log n}}\left(\sum_{i=1}^{n}(1-\frac{i}{n})^{k-1}X_{i:n}-\frac{\theta_k}{k}\right) \qquad (3.34)\\
&=\overset{+}{(-)}\ k\sigma_k \quad \text{a.s.}
\end{aligned}$$

$$n^{1/2}(T_n-\theta_k) \xrightarrow{\text{d}} N(0,k^2\sigma_k^2) \tag{3.35}$$

$$\frac{k}{\sqrt{n}}\left(\sum_{i=1}^{n}(1-\frac{i}{n})^{k-1}X_{i:n}-\frac{\theta_k}{k}\right) \xrightarrow{\text{d}} N(0,k^2\sigma_k^2) \tag{3.36}$$

where

$$\sigma_k^2 = \text{Var}(Y_k)$$

$$
\begin{aligned}
Y_k &= -(1-U)^{k-1}F^{\leftarrow}(U)+\frac{\theta_k}{k}-\int_0^1\left(I_{\{U\le t\}}-t\right)(1-k)(1-t)^{k-2}F^{\leftarrow}(t)dt \\
&= \int_0^1\left(I_{\{U\le t\}}-t\right)(1-t)^{k-1}dF^{\leftarrow}(t)dt.
\end{aligned}
$$

__Remark__ If we replace

$$\min\{X_{i_1},...,X_{i_k}\}$$

by

$$\max\{X_{i_1},...,X_{i_k}\}.$$

Then, similarly we have that

$$
\begin{aligned}
&\frac{1}{\binom{n}{k}}\sum_{(n,k)}\max\{X_{i_1},...,X_{i_k}\} \\
&= \frac{1}{\binom{n}{k}(k-1)!}\left(\sum_{i=1}^n i^{k-1}X_{i:n}+D_1\sum_{i=1}^n i^{k-2}X_{i:n}+\cdots+D_{k-1}\sum_{i=1}^n X_{i:n}\right)
\end{aligned}
$$

where $D_1,\ D_2,\ ...,\ D_{k-1}$ are constants depending on k only. Thus, we can state an analogue of Theorem 3.4 for U-statistic

$$\frac{1}{\binom{n}{k}}\sum_{(n,k)}\max\{X_{i_1},...,X_{i_k}\},\quad n\ge k$$

and L-statistic

$$\frac{k}{n}\sum_{i=1}^n\left(\frac{i}{n}\right)^{k-1}X_{i:n},\quad n\ge 1$$

if we want. Thus, we can not only give an alternative proof of Theorem 1.1 of Gilat and Hill (1992), but also improve their result. For example, under the assumption (3.24), we have that

$$\limsup_{n\to\infty}\frac{k}{n}\left|\sum_{i=1}^n\left(\frac{i}{n}\right)^{k-1}X_{i:n}\right|<\infty\quad \text{a.s.} \tag{3.37}$$

if and only if

$$E\left(|\max\{X_1,\ \ldots,\ X_k\}|\right) < \infty. \tag{3.38}$$

In either case, we have that

$$\lim_{n\to\infty} \frac{k}{n^k} \sum_{i=1}^{n} i^{k-1} X_{i:n} = E\left(\max\{X_1,\ \ldots,\ X_k\}\right) \quad \text{a.s.} \tag{3.39}$$

Example Let $\{X, X_n;\ n \geq 1\}$ be a sequence of i.i.d. random variables with common density function

$$f(x) = \frac{5}{6(-x)^{11/6}} I_{\{x \leq -1\}}.$$

It is easy to check that

$$E(|X|) = \infty.$$

Thus, Theorem 1.1 of Gilat and Hill (1992) can not be applied to

$$\lim_{n\to\infty} \frac{k}{n^k} \sum_{i=1}^{n} i^{k-1} X_{i:n}.$$

However, note that $\frac{3}{4} > \frac{2}{3}$,

$$E(|X|^{3/4}) < \infty$$

and

$$E\left(|\max\{X_1,\ \ldots,\ X_k|\right) < \infty, \quad \text{for all } k \geq 2.$$

So (3.37), (3.38) and (3.39) imply that, for every $k \geq 2$,

$$\begin{aligned} \lim_{n\to\infty} \frac{k}{n^k} \sum_{i=1}^{n} i^{k-1} X_{i:n} &= E\left(\max\{X_1,\ \ldots,\ X_k\}\right) \\ &= \int_{-\infty}^{-1} \frac{5kx}{6(-x)^{(5k/6)+1}} dx \\ &= -\frac{5k}{5k-6} \quad \text{a.s.} \end{aligned}$$

4 The Law of the Logarithm for L-Statistics

Let $\{X,\ X_n;\ n \geq 1\}$ be a sequence of i.i.d. random variables with

$$E(X) = 0 \quad \text{and} \quad E(X^2) = 1 \tag{4.1}$$

and let $J(\cdot)$ be a Lipschitz function of order 1 defined on $[0,1]$. LIL for weighted sums of form $\sum_{i=1}^{n} J(\frac{i}{n})X_i,\ n \geq 1$ established by Tomkins (1975, 1976) implies that

$$\limsup_{n\to\infty} \frac{\sum_{i=1}^{n} J(\frac{i}{n})X_i}{\sqrt{2n\log\log n}} = \left(\int_0^1 J^2(t)dt\right)^{1/2} \quad \text{a.s.} \tag{4.2}$$

By comparing (4.2) with (2.7), we can find that both (4.2) and (2.7) have the same framework except for two differences. One of them is, usually,

$$\int_0^1 J^2(t)dt \neq \text{Var}(Y)$$

where Y is defined as in (2.2). On the other hand, Li, Rao and Wang (1995) show that, for almost all chices $\{c_{i,n};\ 1 \leq i \leq n,\ n \geq 1\}$ of triangular arrays of real numbers with

$$\sup_{1\leq i\leq n, n\geq 1} |c_{i,n}| < \infty,$$

the weighted sums

$$\sum_{i=1}^{n} c_{i,n}X_i, \quad n \geq 1$$

obey the Law of the Logarithm, i.e.,

$$0 < \limsup_{n\to\infty} \frac{|\sum_{i=1}^{n} c_{i,n}X_i|}{\sqrt{2n\log n}} < \infty \quad \text{a.s.} \tag{4.3}$$

It follows that

$$\limsup_{n\to\infty} \frac{|\sum_{i=1}^{n} c_{i,n} X_i|}{\sqrt{2n \log\log n}} = \infty \quad \text{a.s.}$$

In this section, we give a version of (4.3) for L-statistics with no proof. See Li, Rao and Tomkins (1993) for the details of its prrof. Before we state our result, we introduce some more notation. Let $\mathcal{I} = \{(i,n);\ 1 \le i \le n,\ n \ge 1\}$. For a given probability measure $\boldsymbol{\nu}$ on the Borel σ-algebra of $\mathcal{R}$, let $\mathcal{S} = \mathcal{S}(\boldsymbol{\nu}$ denote the support of $\boldsymbol{\nu}$. We will consider only those probability measures for which $\mathcal{S}$ is bounded. Let $P' = \boldsymbol{\nu}^{\mathcal{I}}$ be the product probability measure on the Borel σ-algebra of $\mathcal{S}^{\mathcal{I}}$.

<u>**THeorem 3.1**</u> (A Law of the Logarithm for L-Statistics) Let $\{X, X_n;\ n \ge 1\}$ be a sequence of i.i.d. random variables and $H(\cdot)$ a real valued measureable function defined on $\mathcal{R}$ such that

$$E\left(H^2(X)\right) < \infty.$$

Then, for any given probability measure $\boldsymbol{\nu}$ on the Borel σ-algebra of $\mathcal{R}$ with bounded support $\mathcal{S} = \mathcal{S}(\boldsymbol{\nu})$, there exists a set $\Omega_0 \subset \mathcal{S}^{\mathcal{I}}$ such that

$$P'(\Omega_0) = 1 \tag{4.4}$$

and for any $\{c_{i,n};\ (i,n) \in \mathcal{I}\} \in \Omega_0$,

$$\limsup_{n\to\infty}(\liminf_{n\to\infty}) \frac{\sum_{i=1}^{n} c_{i,n} H(X_{i:n}) - n\mu}{\sqrt{2n \log n}} = \overset{+}{(-)}\, \sigma \quad \text{a.s.} \tag{4.5}$$

where

$$\mu = E(H(X)) \int_{\mathcal{S}} t\boldsymbol{\nu}(dt) \quad \sigma^2 = E\left(H^2(X)\right) \left(\int_{\mathcal{S}} (t - \int_{\mathcal{S}} s\boldsymbol{\nu}(ds))^2 \boldsymbol{\nu}(dt) \right).$$

Remark If $E(H(X)) = 0$ and $0 < E\left(H^2(X)\right) < \infty$, then for almost all choices $\{c_{i,n};\ (i,n) \in \mathcal{I}\}$ of triangular arrays of real numbers with $\sup_{(i,n)\in\mathcal{I}} |c_{i,n}| \leq M < \infty$ for some constant $M > 0$, the L-statistics

$$\sum_{i=1}^{n} c_{i,n} H(X_{i,n}), \quad n \geq 1$$

obey the Law of the Logarithm, i.e.,

$$0 < \limsup_{n\to\infty} \frac{|\sum_{i=1}^{n} c_{i,n} H(X_{i:n})|}{\sqrt{2n \log n}} < \infty \qquad \text{a.s.}$$

It follows that

$$\limsup_{n\to\infty} \frac{|\sum_{i=1}^{n} c_{i,n} H(X_{i:n})|}{\sqrt{2n \log\log n}} = \infty \qquad \text{a.s.}$$

Thus, for almost all choices $\{c_{i,n};\ (i,n) \in \mathcal{I}\}$ of triangular arrays of real numbers with $\sup_{(i,n)\in\mathcal{I}} |c_{i,n}| \leq M < \infty$, the LIL for L-statistics

$$\sum_{i=1}^{n} c_{i,n} H(X_{i,n}), \quad n \geq 1$$

fails.

NOTE

Some of the results of this paper were taken from the technical report of Li, Rao and Tomkins (1993).

References

GILAT, D. and HILL, T. P. (1992). One-sided refinements of the strong law of

large numbers and Glivenko-Cantelli theorem. *Ann. Probab.* **20** 1213-1221.

HELMERS, R. (1977). A strong law of large numbers for linear combinations of order statistics. Report SW50/77, Mathematisch Centrum, Amsterdam, The Netherlands.

KOLMOGOROV, A. N. (1933). Sulla determinazione empirica di una legge di distribuzione. *Giorn. Ist. Ital. Attuari* **4** 83-91.

LEE, A. J. (1990). U-Statistics. *Marcel Dekker, Inc., New York.*

LI, DELI, RAO, M. BHASKARA and TOMKINS, R. J. (1993). The strong law of large numbers and the law of the iterated logarithm for liear combinations of order statistics. *Research Report No. 93-?, Center for Multivariate Analysis, Pennsylvania State University, University Park, Pennsylvania 16802.*

LI, DELI, RAO, M. BHASKARA and WANG, X. C. (1995). On the strong law of large numbers and the law of the logarithm for weighted sums of independent random variables with multidimensional indices. *J. Multivariate Anal.* **52** 181-198.

MASON, D. M. (1982). Some characterizations of strong laws for linear functions of order statistics. *Ann. Probab.* **10** 1051-1057.

MASON, D. and SHORACK, G. (1992). Necessary and sufficient conditions for asymptotic normality of L-statistics. *Ann. Probab.* **20** 1779-1804.

SEN, P. K. (1978). An invariance principle for linear combinations of order statistics. *Z. Wahrscheinlichkeitstheorie verw. Gebiete.* **42** 327-340.

SERFLING, R. J. (1980). Approximation Theorems of Mathematical Statistics, *John Wiley, New York.*

SHORACK, G. R. (1972). Functions of order statistics. *Ann. Math. Statist.* **43** 412-427.

SHORACK, G. R. and WELLNER, J. A. (1986). Empirical Processes with Applications to Statistics. *Wiley, New York.*

STIGLER, S. (1974). Linear functions of statistics with smooth weight functions. *Ann. Statist.* **2** 676-693.

SMIRNOV, N. N. (1944). An approximation to distribution laws of random quantities determined by empirical data. *Uspehi Mat. Nauk.* **10** 179-206.

TOMKINS, R. J. (1975). Iterated logarithm results for weighted averages of martingale difference sequences. *Ann. Probab.* **3** 307-314.

TOMKINS, R. J. (1976). Strong limit theorems for certain arrays of random variables. *Ann. Probab.* **4** 444-452.

VAN ZWET, W. R. (1980). A strong law for linear functions of order statistics. *Ann. Probab.* **8** 986-990.

WELLNER, J. A. (1977a). A Glivenko-Cantelli theorem and strong laws of large numbers for functions of order statistics. *Ann. Statist.* **5** 473-480.

WELLNER, J. A. (1977b). A law of the iterated logarithm for functions of order staistics. *Ann. Statist.* **5** 481-494.